

Weekly report for the Epiwork project
Week: 13-17 July, 2009.

Luis Filipe Lopes

Activities in which I have been involved in the past week:

1- Worked on the catalogue deliverable. Collection of epidemiological scientific articles, to use in the development of the catalogue scheme.

2- Reviewing and rewriting the twitter article submitted to the INForum.

Tasks for the next week:

- Finish reviewing the article and submit new version next Monday (20th of July).
- Work on the catalogue deliverable.

Patrícia Sousa

Activities in which I have been involved in the past week:

1- Reinstallation of CentOS in the "Ochato" server.

2- Installation of all the software needed to run Fedora Commons and Muradora.

3- Development of a PHP script to automatically change the user password at the OpenLdap directory.

4- Attempt to integrate the PHP script with the JSP page.

5- Search and aquisition of bibliography to the study of JSP.

Problems:

- The PHP script doesn't seem to work integrated in the Muradora pages.

Tarefas agendadas para a semana:

Tasks for the next week:

- Continue installing the necessary software in the "Ochato" server.

- Try to troubleshoot the PHP script problem.

- Start studying JSP.

- Add to Muradora the possibility of registering new users, using JSP.

Hugo Ferreira

Activities in which I have been involved in the past week:

1- Installation and configuration of the CentOS in the "Ochato" server.

2- Testing of automatic backups in the "Ochato" server, using the Crontab and Tar commands. Use of the Time command to measure the time taken by the backup process.

3- Schedualling of weekly total backups and daily incremental backups in the "Ochato" server.

4- Total backup of the "Remi" server contents, that was kept at the Iomega directory /media/iom01/store.

Tasks for the next week:

- Will not be present.

João Zamite

Activities in which I have been involved in the past week:

1- Reviewing and rewriting the twitter article submitted to the INForum.

2- Studying the UMLS model.

Tasks for the next week:

- Finish reviewing the article and submit new version next Monday (20th of July).

- Continue working on the UMLS.

Weekly report for the Epiwork project
Week: 20-24 July, 2009.

Luis Filipe Lopes

Activities in which I have been involved in the past week:

1- Finished re-writing the article to be published in INFORUM, and its submission.

2- Worked on the catalogue deliverable report.

Have been analyzing in more detail the Dublin Core documentation and its fields in order

to better understand how it could be adapted for the EM.

Worked on a preliminary version of a UML class diagram, to represent the epidemiological

information that needs to be addressed by the catalogue.

Tasks for the next week:

- Continue working on the catalogue report. Improve the UML scheme.

João Zamite

Activities in which I have been involved in the past week:

- 1- Finished the review on the article to be published in INFORUM and prepared the camera-ready copy in LaTeX.
- 2- Explored the UMLS Ontology with emphasis on preparing a few queries for disease names extraction since it not only provides the scientific names for diseases but also some

common names and acronyms:

2.1 - Focused on three tables in the ontology. MRCONSO which relates each string to a specific CUI(concept ID). MRREL which contains the relations between CUIs and their types(child, parent, broad, narrow, etc.). And MRSTY which relates each concept with a semantic term.

This so far resulted in the following two queries:

```
SELECT m.CUI, m.STR
FROM MRCONSO m, MRREL r, MRSTY s
WHERE m.CUI = r.CUI1
AND m.CUI = s.CUI
AND r.CUI2 LIKE "C0009450"
AND s.STY LIKE "%disease%";
```

This query retrieves the concept ID and strings for every CUI related to Infectious Diseases (C0009450) that is a disease (STY term "%disease%").

Results in 172 CUIs which have 5,563 Strings. Problems: does not retrieve subtypes of diseases or their names (retrieves Influenza but not H1N1).

```
SELECT m.CUI, m.STR
FROM MRCONSO m, MRSTY s,
WHERE m.CUI = s.CUI
AND s.STY LIKE "%disease%";
```

This query retrieves all the CUIs and strings which are diseases. This results in 39,450 CUIs and 127,276 Strings. Problems: retrieves all subtypes of the diseases, but does not relate them, also includes things as confirmed infections, diseases with the addition of other symptoms, which shouldn't be included in our searches.

Tasks for the next week:

- Continue working on disease retrieval from the UMLS ontology. An approach would be to use the first query described above, and then using a second query to drill-down on each disease for "child" terms in MRREL which are also diseases.
- Referencing each of the diseases in the Collector database to the UMLS ontology via a CUI.

Patrícia Sousa

Activities in which I have been involved in the past week:

- Complete the installation of all the software needed to run fedora Commons and Muradora on the Ochato.
- Confirmation that PHP script doesn't work on Muradora environment
- Continue studying JSP
- Solve other server problems of the XLDB group.
- Backup of the Meireles server contents to the IOM01.

Tasks for the next week:

- Starting to program JSP script
- Format the Meireles server
- Install CentOS in the Meireles server
- Installation of all the software needed to run the XLDB group web page.

Weekly report for the Epiwork project

Week: 27-31 July, 2009.

Luis Filipe Lopes

Activities in which I have been involved in the past week:

- 1- Worked on the catalogue deliverable. Altered the UML class diagram.

2- Reading of some epidemiological articles in order to understand how well does the information contained in those articles conform to the data scheme of the UML.

Tasks for the next week:

- Continue reading the epidemiological articles.
- Work on the report text.
- Attend some talks in the Perl meeting here at FCUL.

Patrícia Sousa

Activities in which I have been involved in the past week:

- 1- Buy material to learn more JSP.
- 2- Do tests and practice JSP on the OChato server.
- 3- Integrate this script with Muradora.
- 4- Reinstallation of all software needed to run Fedora Commons and Muradora.

Tasks for the next week:

- Make backup databases from server Meireles
- Formating de Meireles Server
- Installation of CentOS in the Meireles server
- Installation of all the software needed to run web page of group XLDB
- Make a real script to change the password of users on Muradora environment.

João Zamite

Activities in which I have been involved in the past week:

- 1- Continued working on disease retrieval from the UMLS ontology. Prepared a query that retrieve all the child terms in MRREL which are related to each disease retrieved from the queries described on the previous report.
- 2- Referencing of each disease in the Collector database to a UMLS Concept ID has been delayed to next week.

Tasks for the next week:

- Referencing each of the diseases in the Collector database to the UMLS ontology via a CUI.
- Writing a script (Python or Java) to retrieve aliases for each of the diseases in the Collector database, as well as other diseases of relevance based on the queries prepared in the past two weeks.